

# Ontología del patrimonio de Cantabria

Por Francisca Hernández-Carrascal

**Resumen:** *Experiencia de aplicación de ontologías al dominio patrimonio cultural, el proyecto Ontología del patrimonio de Cantabria está desarrollado y financiado por la Fundación Marcelino Botín. Se presentan los antecedentes del proyecto y una descripción del mismo, destacando el portal semántico y las diferentes aplicaciones que se incluirán en él, como elementos visibles. Se hace una revisión de la normativa internacional utilizada, desde las recomendaciones del W3C para la web semántica a normas específicas para la creación de una ontología sobre patrimonio cultural como la norma ISO 21127 y otros modelos de datos relacionados como FRBR y Dublin Core. Se describe la metodología seguida para la integración de datos procedentes de fuentes diversas y heterogéneas y los procesos de conversión realizados para crear una base de datos de conocimientos sobre el patrimonio de Cantabria.*

**Palabras clave:** *Ontologías, Patrimonio Cultural, Integración de conocimientos, ISO 21127, Modelo Conceptual de Referencia del Cidoc, Cidoc-CRM, Requisitos funcionales para registros bibliográficos orientados a objetos, FRBROO*

**Title:** **Cantabria's Heritage Ontology**

**Abstract:** *Description of an ongoing project that aims to model the cultural heritage domain in an ontology, the Cantabria's Heritage Ontology, lead by Fundación Marcelino Botín. The ontology will contain data from different and heterogeneous sources (bibliographic records, finding aids, catalogues, etc.). The project will develop a series of applications that get the most out of the ontology, beginning by a semantic portal. The portal will provide easy and intuitive access to a huge volume of information that comes from heterogeneous sources selected by domain experts. People interested in Cantabria will be able to find with as few clicks as possible, not only the information they are interested in, but also some contextual information that complements the one the user is watching, suggesting new navigation directions, and uncovering hidden relations. The ontology has been designed based on the international standards for cultural models, such as CRM, FRBR or Dublin Core, and W3C recommendations for the semantic web.*

**Keywords:** *Ontologies, Cultural heritage, Knowledge integration, ISO 21127, Cidoc conceptual reference model, Cidoc CRM, Functional requirements for bibliographic records object oriented, FRBROO.*

**Hernández-Carrascal, Francisca.** "Ontología del patrimonio de Cantabria". En: *El profesional de la información*, 2008, enero-febrero, v. 17, n. 1, pp. 92-98.

DOI: 10.3145/epi.2008.ene.11



**Francisca Hernández-Carrascal**, Miembro del Cuerpo Facultativo de Bibliotecas. Ha sido directora del Departamento de Control Bibliográfico de la Biblioteca Nacional (1993-1996), directora de la Hemeroteca Nacional (1996-1999) y directora técnica del Archivo Virtual de la Edad de Plata (1999-2003). Ha participado en diferentes proyectos, nacionales e internacionales, relacionados con ontologías y actualmente es la directora del proyecto Ontología del patrimonio de Cantabria que lidera la Fundación Marcelino Botín.

## Descripción

**EL PROYECTO DE LA ONTOLOGÍA del patrimonio de Cantabria**, liderado por la Fundación Marcelino Botín se inició en 2006 y finalizará en 2009. Cuenta con la colaboración del Departamento de Historia Moderna y Contemporánea de la Universidad de Cantabria que ha formado un equipo de investigación dirigido por Manuel Suárez-Cortina y coordinado por Gonzalo Capellán. Este equipo de 'expertos en el dominio' está actualizando la recopilación bibliográfica base Historia

**de Cantabria: un siglo de historiografía y bibliografía (1900-1994)**<sup>1</sup> editada por Manuel Suárez Cortina<sup>2</sup> sobre la que se está construyendo el edificio de la ontología.

La empresa española *iSOCO* (*Intelligent Software for the Networked Economy*) es la encargada de la construcción e implantación del sistema.

El objetivo del proyecto es aglutinar de una forma estructurada el conocimiento sobre el patrimonio de Cantabria en todos sus múltiples aspectos: desde el arqueológico al industrial, desde el científico-cultu-

ral al etnográfico, pasando por sus actores y períodos históricos. Es un campo muy amplio y heterogéneo y está disponible de forma fragmentada, procede de fuentes como publicaciones oficiales, monografías, artículos, catálogos, inventarios, enciclopedias, bases de datos, sitios web, etc., y se encuentra en diferentes estados de formalización y digitalización, todo lo cual dificulta el acceso a la información y su explotación a través de la Web.

Basándose en las tecnologías de la web semántica, el proyecto está elaborando una base de conoci-

mientos (ontología) procedentes de las diferentes áreas científicas y técnicas relacionadas con el patrimonio (la historia, la geografía, el arte, la gestión del patrimonio, etc.) que han sido, y son, generados por diferentes personas y entidades en el ejercicio de sus actividades (instituciones de la administración pública, universidades, entidades privadas, científicos, escritores, etc.), para ofrecerlos en forma de un portal web del patrimonio de Cantabria.

### **Qué es y para qué sirve una ontología**

Como se sabe, las ontologías definen los conceptos y las relaciones que representan un área de conocimiento. Categorizan las condiciones de uso de los términos y sus posibles relaciones, así como las restricciones en el uso de las mismas. En la práctica las ontologías pueden ser muy complejas<sup>3</sup> (con unos varios miles de términos y condiciones) o muy simples (la descripción de uno o dos conceptos solamente). Así, la Ontología del patrimonio de Cantabria contempla los diferentes tipos de patrimonio tal y como están establecidos en la legislación vigente en Cantabria, los objetos, las piezas, los elementos que componen ese patrimonio; las publicaciones que se han editado a lo largo de la flecha del tiempo para difundir, estudiar, analizar, proteger y gestionar el patrimonio; las instituciones que lo gestionan, sus cometidos, sus actuaciones concretas; las personas que han trabajado, investigado, conservado, restaurado y difundido ese patrimonio. Todo ello exige la definición de las clases apropiadas para expresar entidades temporales como períodos cronológicos, períodos culturales, acciones, etc.; objetos biológicos y objetos realizados por el hombre, etc.

El amplio conjunto de datos relacionados con el patrimonio ordenados semánticamente proporcionarán a cualquier tipo de usuario

una gran potencia de búsqueda y navegación, así como le permitirán proyectar o extrapolar el citado conocimiento a actuaciones de todo tipo, desde la recuperación de un dato concreto a la gestión de un proyecto cultural, o a la promoción de un servicio turístico, y, esto no es lo menos importante, a la realización de un modelo de conocimiento transferible a otros ámbitos territoriales, Comunidades Autónomas o países.

Se quiere también con este sistema de organización de la información contribuir a la potencia y versatilidad de recuperación de la información que pretende la web semántica, en la que los agentes de software harán uso del conocimiento vertido en un entramado de ontologías. Del mismo modo que la web actual ha adquirido su potencia por medio de la creación y transformación de contenidos desde todos los puntos del planeta, la web semántica podrá mostrar su verdadero potencial con un esfuerzo similar de estructuración de información y generación de conocimientos. Los métodos y procedimientos, así como las aplicaciones que se están llevando a cabo en este sentido son muy variados, lo que se refleja, entre otros, en *Semantic Web Education and Outreach Interest Group: Case Studies and Use Cases*<sup>4</sup>, página del W3C destinada a casos de estudio y uso en producción relacionados con web semántica en el ámbito cultural<sup>5</sup>, en la que está incluido el proyecto *Ontología del patrimonio de Cantabria*<sup>6</sup>.

### **Fuentes de información**

Dentro del proceso de construir la ontología está la recopilación del conocimiento e información publicado sobre el patrimonio de Cantabria. Se están siguiendo las técnicas de la bibliografía para la categorización y organización de los diferentes documentos que puedan existir sobre un determinado tema,

desde un nivel superior en el que se encuentran las bibliografías de bibliografías, hasta el nivel más profundo en el que se situarían los documentos primarios, bien entendido que en la sociedad actual esto supone la identificación y registro de documentos que pueden encontrarse en múltiples soportes, incluido el electrónico, o estar organizados en bases de datos. Así, en un primer nivel se dispondría el conocimiento publicado en el que entrarían todas aquellas fuentes (bio-bibliografías, bibliografías especializadas, bibliografías de publicaciones periódicas, biografías, diccionarios y enciclopedias, recopilaciones de fuentes documentales para la historia de Cantabria, directorios, catálogos y bases de datos bibliográficas, etc.) que en su conjunto recopilan todo lo publicado sobre Cantabria. Esto equivale a decir, lo que será muy significativo, todo lo estudiado sobre esta Comunidad. Por ello, y como ya se ha dicho, la elección de la obra *Historia de Cantabria: un siglo de historiografía y bibliografía (1900-1994)* no es un mero punto de inicio sino los cimientos de la ontología.

En un segundo nivel debemos considerar las descripciones de las unidades (materiales bibliográficos, documentos de archivo, bienes de interés cultural muebles e inmuebles, etc.). En este punto los catálogos de bibliotecas y museos; los registros, inventarios y catálogos de bienes de patrimonio; las guías, inventarios e instrumentos de descripción de archivos, etc., serán las fuentes fundamentales de documentación. No deben olvidarse en este punto las estructuras de metadatos normalizadas en las que se transmite esta información (formato Ibermarc, MARC 21, EAD, EAG), puesto que se tendrá que trabajar con ellas. Ni, por supuesto debe dejarse de lado, que muchas de estas fuentes se hallan en bases de datos no normalizadas, en formato texto sin estructura alguna o incluso ni siquiera existen

en forma digital. Esta misma situación es la que podemos encontrar en el tercer nivel, el conocimiento sobre la gestión del patrimonio. En este apartado podemos situar la legislación sobre el patrimonio, los diferentes tipos de administraciones que actúan sobre él y sus instrumentos de difusión y actuación (programas, proyectos, etc.). Por último, en un nivel de conocimiento general, se han identificado las fuentes de información (y sus correspondientes estructuras de información) que permitirán obtener datos concretos sobre lugares geográficos, personas, cronologías, instituciones históricas y actuales, etc.

### Aplicaciones

Construidas sobre la ontología se dispondrá de una serie de aplicaciones que se irán finalizando en el portal a lo largo del proyecto y que permitirán una explotación específica de los datos. Un buscador semántico permitirá búsquedas que habitualmente no pueden realizarse como “libros impresos entre el 1907-1939, escritos por autores nacidos en Santander” y mediante un mapa interactivo se podrán geoposicionar todos los datos del patrimonio de Cantabria, así como establecer rutas entre diferentes puntos. A partir del conocimiento acumulado podrán crearse recursos didácticos digitales, construyendo automáticamente cursos y exámenes sobre los distintos movimientos artísticos, yacimientos, espacios naturales o conjuntos históricos. De los datos obtenidos de proyectos sobre el patrimonio se construirá una herramienta que analice las acciones ya realizadas, y extrapole y realice estimaciones de propuestas concretas de actuación. Igualmente, una aplicación permitirá configurar rutas turísticas según el interés de los usuarios; completar aplicaciones de navegación de vehículos (GPS) con información de interés sobre la zona visitada o aplicaciones automáticas de guías turísticas



Figura 1



Figura 2

para completar la información sobre puntos visitados (transferibles a dispositivos móviles). Por último se generará una Wikipedia semántica con los términos de la ontología con funcionalidades que permitan la actualización de datos, la aportación de comentarios, informaciones, documentos, etc.

### Normativa

No es el objetivo de este artículo entrar en el detalle de los lenguajes de marcado necesarios para expresar ontologías. Por supuesto, el proyecto se vale de las recomendaciones del W3C para la web semántica<sup>7</sup>, y está utilizando RDF y OWL. Inicialmente RDF fue creado como



lenguaje de XML para la expresión de relaciones semánticas y OWL como lenguaje específico para la construcción y transmisión de ontologías en la web. Sin embargo, hay que considerar que el software propio de la web semántica (piezas de software, aplicaciones, librerías, etc.) está haciendo uso de ambos por su capacidad para ser reutilizado en diferentes entornos y aplicaciones. Hay que mencionar que dentro de OWL<sup>8</sup> existen diferentes sabores como OWL Lite, OWL DL, o OWL Full y que algunas aplicaciones específicas como los razonadores funcionan sólo con uno de los sabores mencionados.

Ha sido preciso definir una estructura común para las informaciones y datos mencionados para lo cual, se tomaron los dos modelos más sólidos que existen: los *Functional requirements for bibliographic records* de IFLA y el *Cidoc conceptual reference model (CRM)*<sup>9</sup>. Éste fue elaborado por el *Comité Internacional de Documentación (Cidoc)* del *Consejo Internacional de Museos*, que se transformó en 2006 en la norma *ISO 21127 Information and documentation-A reference ontology for the interchange of cultural heritage information*. La norma ISO 21127 proporciona en una estructura formal las definiciones necesarias para describir los conceptos y relaciones usadas en la documentación sobre patrimonio cultural. Constituye un marco semántico común y extensible al que puede ser convertido cualquier información sobre patrimonio cultural y es una guía para el modelado conceptual de las entidades (*entities*) que forman el patrimonio cultural.

Se ha considerado imprescindible seguir esta norma, primero para asegurar la interoperabilidad de esta ontología con otras análogas, y segundo para poder transferir y ser objeto de transferencia de conocimientos desarrollados por terceros,

incluido el software y los aplicativos elaborados conforme a la referida norma ISO. En otras palabras, su uso permite la transformación de diferentes fuentes en una ontología común e interoperable, lo que es sustancial a la web semántica.

En tanto que norma internacional, permite la interconexión de diferentes ontologías del patrimonio creadas en torno a diferentes ámbitos temáticos, geográficos o administrativos. Es perfectamente razonable pensar una ontología global sobre el patrimonio formada por la suma de ontologías parciales de ámbitos más restringidos. Este es por otra parte el objetivo que anima la web semántica en sí.

En la Figura 3 puede verse la estructura de CRM con las clases y relaciones de más alto nivel del modelo<sup>10</sup>.

La norma ISO 21127 reconoce la necesidad de desarrollar determinados aspectos de la ontología y prefigura el modo en el que esto puede llevarse a cabo. Indudablemente en el marco de la Ontología del patrimonio de Cantabria ha sido necesario concretar nuevas clases, siempre incluidas dentro de alguna de las definidas en la norma, para

expresar, por ejemplo, relaciones familiares o profesionales entre personas; ocupaciones, profesiones y relaciones de personas con entidades.

Por otro lado el modelo conceptual del CRM ha tenido desde sus inicios una fuerte influencia museística. Como ya se ha comentado están muy detalladas determinadas acciones, pero no así otras, como los aspectos bibliográficos. Obviamente el mejor análisis de la información bibliográfica es el establecido en los *Requisitos funcionales para registros bibliográficos (FRBR)*<sup>11</sup> y en consecuencia, en 2003 se formó un grupo de trabajo, *Working Group on FRBR/CRM Dialogue*<sup>12</sup>, con el objetivo de una diseñar una formulación orientada a objetos de FRBR que actuara como extensión de CRM. En mayo de 2007, aunque ya existía alguna versión anterior, se publicó el borrador de la versión 0.8.1c de *FRBR<sub>OO</sub>* (denominación que la distingue de la versión entidad-relación o *FRBR<sub>ER</sub>*). Hay que mencionar que la incorporación de esta extensión es conforme con la norma ISO 21127 puesto que las clases específicas definidas en FRBR (p. e., F1.Work, F2.Expression, F3.Manifestation\_Product\_Type o

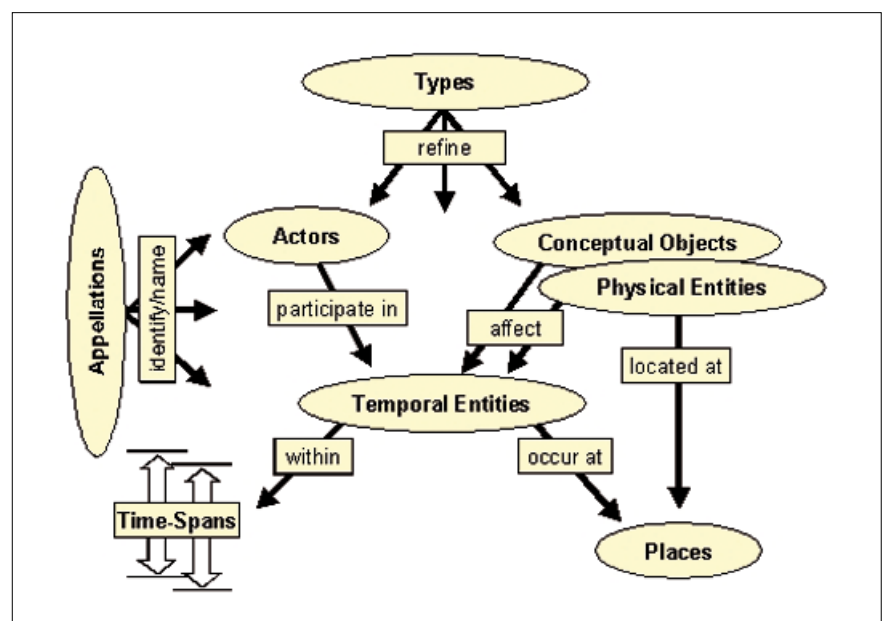


Figura 3

F5.Item) son subclases de clases ya definidas en *CRM* o bien se trata de clases equivalentes. La Ontología del patrimonio de Cantabria ha incorporado esta extensión de *CRM* a la ontología cubriendo así el modelado de la información bibliográfica.

En la concepción del *CRM* y, por tanto, de la norma *ISO 21127*, estaba la posibilidad de convertir datos procedentes de archivos, bibliotecas y museos. Así se han realizado propuestas, más bien teóricas, de transformación desde *Dublin Core*, *Encoded Archival Description*, *MDA Spectrum*, etc. Esto es evidente puesto que el conocimiento de un dominio se obtiene tanto del modelado o estructuración de ese dominio (ontología) como de los datos concretos o instancias que se incorporan a la ontología y esos datos concretos se encuentran en registros con esas estructuras de información. Es más, aunque se trate de procesos completamente diferentes intelectual y técnicamente, son insolubles; una ontología está formada tanto por la estructura de la información, el modelado del dominio, como por los datos propios de ese dominio. El trabajo de armonización entre *CRM* y *FRBR*, no disponible al inicio del proyecto y que todavía se encuentra en estado de borrador, no incluye la definición de los mapeos de información codificada. O mejor dicho, pueden encontrarse propuestas de definición de conversiones (desde *EAD*, *DC*, etc.) previas a la fusión *CRM/FRBR*, que ha sido necesario rehacer en el proyecto, puesto que la inclusión de algunas clases de *FRBR* afecta no sólo a los materiales bibliográficos impresos, sino también a los manuscritos y documentos de archivo.

*Dublin Core* (*DC*) representa un caso especial dentro de la Ontología del patrimonio de Cantabria porque es una de las estructuras de información en la que están defini-

dos numerosos registros de interés para el patrimonio de Cantabria (pensemos que es la norma de codificación de registros que sustenta los repositorios *OAI*), pero también se utiliza para la gestión de la ontología describiendo cada uno de sus elementos (clases, propiedades e instancias) como recursos informativos. Así, cada instancia tiene asociado a través de *DC* su autor, su título, su fecha de creación, la fuente de procedencia, etc., lo cual permite realizar controles típicos como el seguimiento de un proceso o un programa de conversión o la procedencia de un determinado dato.

## Avances

**Adquisición.** Para la integración masiva de datos en la ontología se ha implementado una serie de herramientas de adquisición de datos, cada una de ellas relacionadas con estructuras típicas de información y compuestas por un conjunto específico de reglas según el tipo de fuente. Así, puede importarse información a partir de sitios web semiestructurados (p. e., tablas), de bases de datos *ODBC* o *JDBC*, de registros en formato *Ibermarc* o *MARC 21* (registros bibliográficos, de autoridades y de fondos y localizaciones), de registros *Dublin Core*, y a partir de ficheros *EAD* o *EAG*<sup>13</sup>. En la actualidad la Ontología del patrimonio de Cantabria está formada por 153 clases y 410 propiedades, así como por más de 110.000 instancias procedentes de registros bibliográficos *MARC 21*, registros *Dublin Core*, y otras fuentes como bases de datos relacionales, sitios web, etc.

**Registros poco informados.** Hay que mencionar en este punto algunas cuestiones importantes para la gestión de una ontología que no son realmente nuevas porque emergen en cualquier proceso de conversión, como son la calidad y granularidad de los registros, una excesiva flexibilidad en el uso de los

formatos de codificación y una estructuración muy débil en las bases de datos. Por ejemplo, la tendencia a elaborar registros bibliográficos cada vez más reducidos (también en tiempo y coste) hace que estos pierdan información que puede ser sustancial y al menos debería considerarse que uno de los usos de estos registros va más allá de la gestión y control de los procesos de circulación en bibliotecas y que pueden ser reutilizados en otros sistemas de información. Esperemos que el nuevo código de catalogación internacional<sup>14</sup> tenga en cuenta estos usos.

**Duplicados.** Como puede suponerse, al utilizar diversas fuentes de información es frecuente la aparición de instancias duplicadas, para lo que se ha diseñado un programa específico que detecta homonimias y sinonimias, que pueda ser aplicable en los distintos estados de evolución de la ontología.

**Operaciones con fechas.** Algunas de las aplicaciones diseñadas deberán realizar operaciones (sucedió al mismo tiempo, es anterior a, es posterior a, etc.) para lo que se ha desarrollado un programa para la expresión de fechas en un único formato al que pueden trasladarse fechas aproximadas o indeterminadas del tipo 'finales del siglo XX', 'a mediados de la década de los años 60', o bien fechas anteriores o posteriores a Cristo, así como distintas expresiones de indeterminación como 188?. Indudablemente, a lo largo del proyecto se irá completando la definición de diferentes reglas<sup>15</sup>, tanto de análisis de la ontología para la detección de homonimias y sinonimias o para la detección de relaciones no explícitas, como para la visualización, exportación o explotación de los datos

**Visualización.** En cuanto a la visualización de información en la web se ha establecido a través de una ontología de publicación una serie de entidades complejas bási-

cas (periodos cronológicos, periodos culturales, personas, entidades, lugares, temas, y unidades de patrimonio) para las cuales se han definido los elementos básicos de visualización (nombre, fecha de nacimiento y muerte, profesión, actividades, obras, etc.), las preguntas en Sparql<sup>16</sup> que permiten la obtención de datos de clases y propiedades próximas (p. e., los autores cántabros se forman por medio de una 'query' que extrae todas las instancias de la clase persona que han sido o son escritores y que han nacido en Cantabria), así como los elementos de navegación en la web. Es preciso decir que si bien todas las clases y propiedades de la ontología son necesarias para estructurar el conocimiento, no todas deben ser directamente visibles en una interfaz web en aras de una mejor usabilidad, comprensión y navegabilidad. De todas formas, el portal web cuenta con una visualización, destinada a usuarios concretos, que permite ver y navegar por cualquiera de los elementos de la ontología. Entre paréntesis, no puede dejarse de mencionar que hace unos años se consideraba absurdo que la información bibliográfica pudiera visualizarse en formato MARC mientras que ahora, todas las grandes bases de datos bibliográficas ofrecen esa posibilidad.

### Edición y gestión

Entre las dificultades más importantes encontradas está el estado de desarrollo de los editores de ontologías, principalmente destinados a la creación y edición de la estructura de información (las clases, propiedades y restricciones) de una ontología, pero no tanto a la gestión y manipulación de gran cantidad de instancias. Los desarrollos ya efectuados en la visualización de instancias en el portal han superado en parte esta dificultad, pero quedan algunas cosas por mejorar en herramientas de base como *Protégé* o *SemanticWorks*).

### Conclusiones

Conviene insistir en el hecho enormemente significativo de que el CRM está ya caracterizado y normalizado en la ISO 21127 lo cual da a todo este trabajo una estabilidad sumamente grande. Es decir, que es posible, dentro de una estrategia de transferencia tecnológica, crear una metaestructura fruto del análisis FRBR/CRM y del mapeo de estructuras normalizadas como MARC 21, y tampoco sería particularmente costoso el esfuerzo que habría que realizar, para trasladar con facilidad este intangible –sumamente laborioso desde el punto de vista de su realización– a otro territorio, a otra Comunidad Autónoma o a otro país, a un estilo arquitectónico, a una escuela literaria o a la recuperación de un tipo concreto de patrimonio. Se ha logrado formalizar e implementar en un sistema automatizado el conocimiento no ya de los expertos que participan en el proyecto, es decir la *Fundación Marcelino Botín*, la *Universidad de Cantabria* y la empresa *iSOCO*, sino también el de los distintos expertos que han elaborado y formalizado tanto el FRBR como el CRM (por no hablar del formato MARC).

Una de las pruebas para poner de manifiesto la utilidad de todo este trabajo es su tratamiento de información real. En este caso se ha procedido a llevar a cabo pruebas con diferentes catálogos impresos, como por ejemplo el Catálogo de publicaciones periódicas de la *Biblioteca Municipal de Santander* o el *Catálogo de Manuscritos Musicales de Cantabria*.

Naturalmente, la vinculación de este tipo de registros con las grabaciones de un grupo de música gregoriana de los citados manuscritos litúrgicos mostrará la enorme posibilidad informativa de un procedimiento de este tipo y de la metodología ya desarrollada. En este sentido serán muy importantes los mecanismos de colaboración que se

establezcan con diferentes entidades, tanto para la transformación de los datos, estructurados o no, como para la elaboración de programas complementarios conjuntos de digitalización y conversión.

**Ejemplo.** Si bien en la entrada Cabezón de la Sal de la *Gran Enciclopedia de Cantabria* no aparece la relación de escritores nacidos en esta localidad, como tenemos a lo largo de la *Enciclopedia* múltiples entradas que hacen referencia a autores montañeses nacidos en este pueblo, tras crearse las correspondientes instancias podrán suministrarnos de forma inmediata una respuesta a esa cuestión. Pero además también sabremos la música compuesta en torno a esta ciudad, los empresarios, científicos, artistas, políticos... nacidos en ella o que en ella llevaron a cabo su trabajo más relevante, y todo ello establecido a través de la flecha del tiempo. Tendremos en esa misma línea cronológica los procesos comerciales, industriales y más importantes relacionados con Cabezón así como el patrimonio más destacado. Como además pueden establecerse conversiones desde otras fuentes (p. e., desde el *Nomenclator* del INE) y bases de datos geográficas e históricas tendremos, a través del conocimiento estructurado en el sistema, sus representaciones gráficas, cartográficas, sus vías de comunicación tanto actuales como históricas y sus comercios, actividades y vinculaciones con la realidad sociológica que gira en torno a esa entrada. Quien dice un topónimo dice un autor o una obra de arte, o una familia, o un género, o una especie de la fauna y de la flora, o un monumento. Lógicamente será necesario ir ajustando estos procesos a todas las fuentes de información (algunas inéditas como la *Bibliografía de la provincia de Santander*, de Leguina) como las *Guías*, de Coll, los trabajos de erudición local, los nuevos planteamientos técnicos de las universidades, en particular de



la de Cantabria, o las actuaciones históricas y actuales como las de la Comunidad Autónoma.

*"In the Semantic Web, it is not the Semantic which is new, it is the Web which is new"* Chris Welty, IBM

## Notas

1. Suárez-Cortina, Manuel, ed. *Historia de Cantabria: un siglo de historiografía y bibliografía (1900-1994)*. Santander: Fundación Marcelino Botín, 1995.

2. En *Las bibliografías regionales y locales españolas: (evolución histórica y situación actual)*. Madrid: Ollero y Ramos, 2003, Juan Delgado-Casado dice: "no es una historia sino una bibliografía sobre la región que incluye casi 8.000 referencias, centrada sobre todo en textos históricos pero sin evitar la descripción de obras de todo tipo y especialmente relacionadas con el arte y los monumentos... Me parece que es el reperto-

rio fundamental sobre esta región y una obra de importancia indiscutible.

3. The Gene Ontology.  
<http://www.geneontology.org>

4. <http://www.w3.org/2001/sw/sweo/public/Use-Cases>

5. El 12 de noviembre de 2007 se convocó en Corea el *First international workshop on cultural heritage on the semantic web*, en el marco de la 6<sup>th</sup> International semantic web conference.

6. <http://www.w3.org/2001/sw/sweo/public/Use-Cases/FoundationBotin>

7. <http://www.w3.org/2004/01/sws-pressrelease.html.en>

8. <http://www.w3.org/TR/2004/REC-owl-features-20040210/#s4>

9. <http://cidoc.ics.forth.gr>

10. Tomada de: Crofts, Nick; Doerr, Martin; Gill, Tony. *The Cidoc conceptual reference model: a standard for communicating cultural contents*. February 2003.  
<http://www.cultivate-int.org/issue9/chios>

11. <http://travesia.mcu.es/documentos/requisitos.pdf>

12. [http://www.ifla.org/VII/s13/wgfrbr/FRBR-CRMdialogue\\_wg.htm](http://www.ifla.org/VII/s13/wgfrbr/FRBR-CRMdialogue_wg.htm)

13. EAG (Encoded archival guide) es una DTD para la elaboración del Censo-Guía de los archivos españoles.

[http://aer.mcu.es/sgae/jsp/censo\\_guia/seg\\_nivel/presentacion\\_proyecto.html](http://aer.mcu.es/sgae/jsp/censo_guia/seg_nivel/presentacion_proyecto.html)

14. Resource Description and Access (RDA).  
<http://www.collectionscanada.ca/jsc/rda.html>

15. <http://www.w3.org/2005/rules/>

16. El Rule Interchange Format Working Group del W3C trabaja actualmente en una definición de un lenguaje común de reglas, su utilización e intercambio entre sistemas, y su aplicación a OWL y RDF.  
<http://www.w3.org/TR/rdf-sparql-query/>

**Francisca Hernández-Carrascal,**  
Fundación Marcelino Botín  
[hdz@fundacionmbotin.org](mailto:hdz@fundacionmbotin.org)



ANUARIO

**Think**  
**EPI**

## Artículos en la sección CONTENIDOS

### Pero, ¿qué es Google?

Francisco Tosete

### Web y buscadores: ¿una medida de la realidad?

Isidro F. Aguillo

### Función de los portales temáticos en la era de la búsqueda posmoderna

Ernest Abadal y Lluís Codina

### OCLC: 21.985 miembros ¿cuántos españoles?

Adela d'Alòs-Moner y Eva Méndez

### Prensa digital en 2007

Javier Guallar

Más información en <http://www.thinkepi.net/repositorio>

Ya puedes pasarnos tu pedido: [epi@sarenet.es](mailto:epi@sarenet.es)

Anuario ThinkEPI 2008

89,42 € + IVA = 93 €

Anuario ThinkEPI 2007 + 2008

115,39 € + IVA = 120 €

Anuario ThinkEPI

<http://www.thinkepi.net>